**Evaluation strategies for
library/information systems**

**Professor Tom Wilson
University of Sheffield**

**Introduction**

The customary view of library/information system evaluation is coloured by the great deal of attention given to the evaluation of information retrieval systems over the past 25 years. Such terms as 'recall' and 'precision' have entered the professional vocabulary and, even if their use Is not fully understood, they are used in all kinds of contexts - some of them Inappropriate. The dlfficulties associated with evaluation of this kind are well known: there is the difficulty, for example, of determining what is 'relevant' and, indeed, what is 'relevance'. There is the problem of determining suitable measures of relevance, under whatever definition we use, and so on.

In spite of the problems, however, the idea of the evaluation of IR systems is a very powerful idea which has affected librarians' willingness to think of evaluation as a desirable and necessary function to perform In the management of library and information systems. All too often, however, the IR model of evaluation is used as a criterion by which the choice of method of evaluation is determined, with its emphasis upon measurement and trade-offs between sets of variables. Ideally, of course, it is desirable to have methods of evaluation which have these characteristics, but all too often in library systems it is not possible to identlfy such methods.

The reason that IR evaluation succeeds as far as it does is that the objectives of information retrieval systems can be set down in a very explicit manner. One can say, for example, that the ideal IR system Is one which will deliver all the 'relevant' documents from a collection to the user and no 'irrelevant' documents. This ability to identify goals and objectives is an essential prerequisite of all evaluation, because only then can we think about criteria which indicate when an objective has been met, and only then can we think of whether or not it is possible to 'measure' these criteria in a quantitative sense.

**Why do we wish to evaluate services?**

The idea of evaluation seems to be one which has come into the professional consciousness only in very recent years. True, in an earlier time, in the 1950s in the USA, the idea of the 'library survey' developed, and that had in it the seeds of evaluation (indeed, the library survey concept goes back at least to the 1920s in the USA). In more recent times, however, the impetus to evaluate has come from the need to justify budgets more explicitly than ever before. All service functions in all organizations are being reviewed in terms of their necessity for the aims of the organization and libraries and information systems are no exception. In educational institutions faced with budget cuts the same phenomenon occurs, and in local government the public library is subject to the same pressures.

The consequence of this is that the idea of ***cost*** has come to be associated with evaluation and there has been, perhaps, an over-emphasis on costs, at the expense of justifying services on the grounds of usefulness to the library user. The emphasis on money, a useful quantitative measure, has also led, in my opinion, to a down-grading of other criteria for assessing the benefits of library or information services, and I believe that more attention ought to be given to those criteria.

**What can we evaluate?**

The answer to the question "What can we evaluate?" is very simple: any aspect of organizational functioning can be subject to evaluation. Thus, we can evaluate:

* the way the management structure functions;
* internal operations relating to information materials, such as cataloguing and classification, indexing, etc.;
* library/information services to users:
* new programmes of service delivery;
* new possibilities for technological support to services;
* alternative possibilities for doing anything;
* the functioning of a total system prior to planning change;
* etc., etc.

In other words, if a librarian asks, 'Is it possible to evaluate X?", the answer is always 'Yes' because in principle it is possible to evaluate anything, the only problems are that we don't always know how to do it, or it may cost too much, or the methods available may not produce the kind of information the librarian had in mind when s/he asked the question.

**According to what criteria?**

The next question, and one which has a number of Implications, is, 'According to what criteria can we evaluate X?' The following criteria may not exhaust the possibilities, but they will be enough to be going on with:

* 'success',
* efficiency,
* effectiveness,
* benefits, and
* costs, which may be evaluated independently, or in association with any of the above.

One point to bear in mind is that all of these, except the last, may involve either quantitative or qualitative types of assessment, or a combination of both.

**'Success'**

The idea of 'success' as a criterion for evaluation is clearly associated with new programmes, or other kinds of innovation. Equally clearly, what counts as a 'successful' project may be a matter of subjective perception. One person's 'modest success' Is another person's 'failure' - politicians, of course, are adept at turning calamities into 'modest successes', and perhaps chief librarians ought to be equally adept.

We can move away from complete subjectivity only when we can define the criteria for 'success' and, clearly, this will depend on the objectives *of* the new programme or other innovation. For example, we may introduce, in a public library, a new lending service for video-cassettes: is it our aim:

* to fulfil some philosophically-based aim to provide access to information materials of all kinds?
* to earn additional income by charging for the service?
* to attract users to the library service who would not otherwise be users?

Our criteria for success will vary depending which of these aims we have: in the first case, simply to provide the material is to fulfil the aim and, therefore, one has been 'successful'. In the second case we can measure 'success' by determining whether or not we have made a profit: if the income from loans is higher than the annual cost of establishing and maintaining the collection, we can say that we have been successful. The third case is much more difficult: It requires us to keep detailed records of users and to discover by means of surveys whether there is a higher proportion of former non-users of any aspect of library service as a result. We cannot finish there, however: how many new users must we attract to be satisfied that we are 'successful'? Is it enough that they use the video lending service, or do we expect them to use other services also? If the latter, do we have any expectations of frequency of use that would cause us to say whether or not the service has been successful? And so on - the more complex the reasons for doing anything, the more complex the questions become, and the more time-consuming the process of evaluation.

**Efficiency**

We say that something is 'efficient' when it performs with the minimum use of whatever resource is necessary for it to function. Thus, if we are assessing the efficiency of an Internal-combustion engine we will be concerned with how much fuel it uses, relative to other engines of equivalent power. Athletes' efficiency can be measured by their capacity to take in oxygen and convert it to energy.

The efficiency of library or information systems, therefore, must be measured according to the consumption of their resources - people, materials, and money. Such questions as, 'Can we provide the same or better level of reference service with fewer people if we install on-line search terminals?' are likely to arise - particularly if a trade-off between people and capital expenditure is a necessary part of the case to be made. Or, the question may be asked, 'Is it more efficient to buy our cataloguing from a central source, or to do it ourselves?' The assumption, in this second question, of course, is that either alternative is going to provide equally effective catalogues.

'Cost-efficiency' has already been touched upon in one of the above questions - it is concerned with the question 'Can we do X at lower cost by doing it some other way than we do now?' Clearly, to carry out cost-efficiency exercises we need to know the cost of the alternatives and while this may be easy to discover for a system which is already operating, it may be more difficult to discover the real operating costs of a hypothetical alternative.

**Effectiveness**

The effectiveness of any aspect of library operations or service requires a judgement at how well the system is performing relative to its objectives. Lancaster notes that:

'Effectiveness must be measured in terms of how well a service satisfies the demands placed upon it by its users.' (Lancaster, 1977: 1)

but, of course, the effectiveness of aspects of library operations other than public services may be the subject of effectiveness evaluation.

Again, the crucial element in evaluating effectiveness, is a consideration of objectives. For example, if we are assessing the effectiveness of a cataloguing department do we make the criterion the technical one of maintaining an adequate throughput of materials in terms of the size of the backlog, or is the criterion one related to how well the resulting catalogue serves the needs of readers?

Obviously, it is easier to evaluate effectiveness if the criteria are capable of quantitative measurement - it is easier (and cheaper) to count the size of the backlog than it is to relate the quality of catalogues to the needs of users. This ease with which quantitative data can (generally) be collected might be said to have deterred librarians from trying to discover the effectiveness of their systems in terms relating to users, which may be more meaningful.

Cost-effectiveness is concerned with the question, 'Given the level at which the system is performing, can we maintain an equal level of performance at a lower cost?' Again, the costs of the existing system may be discoverable, although this is not necessarily an easy matter, whereas the discovery of the costs of alternative methods of doing things may be very difficult to discover. All changes In organizations have unforeseen consequences, and one of the consequences may well be a higher cost than had been anticipated.

**Benefits**

The benefits of libraries must relate to the communities they serve. We are concerned either with the value of a library (or library service) to the community as a whole (whether the community is a town, an educational Institution, or a firm) or to an individual user. This is the most difficult question for librarians to find an answer to, particularly when the issue of cost is added as in cost-benefit evaluation.

I would argue, however, that simply because it is difficult to attach money values to benefits, this is not an excuse to avoid trying to evaluate benefits at all. After all, when we use the term 'values' In other contexts we mean not money values but philosophical or ethical or moral values - and none of these has money values attached to it as a matter of course.

In other words, I argue that a well-founded qualitative evaluation of benefit can be very persuasive if it supports 'values' of these other kinds held by those responsible for making decisions about resources. I suspect, for example, that a study of the benefits of public libraries in terms of services to the unemployed would find a more sympathetic ear in Sweden than it would in Mrs. Thatcher's Britain.

**By what research strategies?**

If we divide research methods into the two categories "quantitative" and "qualitative" (recognizing that this is a crude distinction and that these terms are the polar extremes of a continuum of methods) the following typology of methods results:

* quantitative
	+ data collection by constant monitoring
	+ data collection by ad hoc surveys of activities or costs
	+ ad hoc surveys to solicit 'hard' data from users
* qualitative
	+ ad hoc surveys to obtain information/opinions from users

Looking a little closer at the typology above, what do we mean by the terms and what skills need to be brought to bear?

* ***monitoring***: collecting data as things occur. The most common type ot monitoring is usually done in relation to the simplest functions of libraries. For example, lending statistics can be collected easily, particularly if they are generated by the computer issue system} data on reference questions are more difficult to collect partly because of the difficulty of identifying generally agreed and comparable definitions of different types of enquiry, but attempts are often made simply to record the crude number of enquiries. Monitoring the 'in-library' use of books and journals is also sometimes attempted by recording the numbers of items left on reading tables. One may also monitor certain kinds of activity not on a continual basis but just occasionally and often by observation, as when we monitor reading room use by taking regular counts throughout the year. The skills needed for any of these activities are pretty rudimentary, the basic skill is that of counting. As a general rule it can be said that the more interesting the activity, the more difficult it is to monitor!
* ***surveys***; are generally of only two kinds - those that use a 'self -completed' questionnaire (sometimes called a 'mail' questionnaire) and those that involve interviewing. The subject is far too large to deal with here and the skills involved range from a knowledge of sampling, through questionnaire design (Including the design of questions, which is not as easy as it might seem), and interviewing skills, to the computer analysis of data. Whereas monitoring may be done by any librarian with a grain of common sense, surveys require at least the consultancy services of a survey researcher.

'Action research' is another mode of research which has a place in evaluation strategies. This method of research is designed to bring about change by identifying problems collaboratively, collecting information to provide the basis for planning action, initiating change, monitoring the results, and repeating the cycle as necessary.

Action research, in other words, can be seen as evaluated innovation and both quantitative and qualitative methods of collecting data, information and opinions may be used in the various data-collection and monitoring phases. Action research is a type of research to use if what you want to do is to change a situation rather than study it.

**Conclusion**

My conclusion is a very simple one - only the librarian or information worker can decide what s/he needs to evaluate. Usually the call for evaluation is the result of some problem occurring in a system, or the result of the librarian being under pressure to justify expenditure on staff, or materials, or other aspects of services, or proposed changes to services. Evaluation for the sake of academic curiosity seems to me to be out of place in organizations which are seeking to fulfil valuable functions for their communities.

Thus, *why* you evaluate is a question you must decide, as is *what* you should evaluate. *How* to evaluate a library service or some library operation is something about which you might need some assistance, and if survey research methods are to be used in the evaluation it is highly desirable to get help from someone skilled in the use of these methods.

There is no *one* way to carry out evaluation - it all depends upon your objectives, the problems they give rise to, and the kinds of information you need in order to make decisions about systems

**Suggested reading**

* Cronin, B. (1982) 'Taking the measure of service'. *Aslib Proceedings*, **34**. 273-294
* De Prospo, E.R., Altman, E. and Beasley, K.E. (1973) *Performance measures for public libraries.* Chicago: ALA.
* King, D.W. and Bryant, E.C. (1971) *The evaluation of information services and products.* Washington, DC: Information Resources Press.
* Lancaster, P.W. (1977) *The measurement and evaluation of library services*. Washington, DC: lnformation Resources Press.
* Lancaster, P.W. and Cleverdon, C.W. eds. (1977) *Evaluation and scietlfic management of library and information services.* Leyden: Noordhoff.
* Orr, R.H. (1973) 'Measuring the goodness of library services: a general framework for considering quantitative measurements'. *Journal of Documentation*. **23**, 315-332
* Wills, G. and Oldman, C. 'An examination of cost/benefit approaches to the evaluation of library and information services', in: Lancaster and Cleverdon (1977) *op cit.* 165-184

On internal evidence this was a paper presented at a meeting in Sweden some time in the mid-1980s, but not having recorded unpublished conference papers I have no idea exactly where or when. If anyone recognizes it and remembers the occasion, I would be most interested to hear from them!